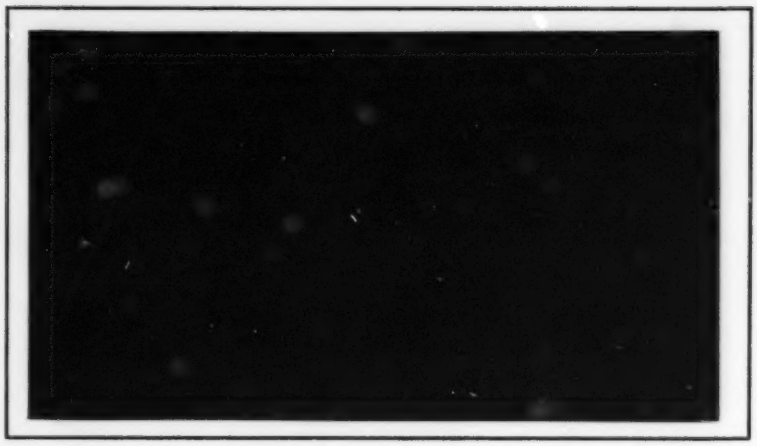

CENTRE FOR
SYSTEMS
SCIENCE

LCCR

Simon Fraser University, Burnaby, B.C., Canada V5A 1S6 / (604) 291-4704

Amplifier Linearization Using A Digital Predistorter with Fast Adaptation and Low Memory Requirements

James K. Cavers

CSS/LCCR TR 90-07

April 1990

92-05789/

AMPLIFIER LINEARIZATION USING A DIGITAL PREDISTORTER WITH FAST ADAPTATION AND LOW MEMORY REQUIREMENTS

by

James K. Cavers

School of Engineering Science

Simon Fraser University

Burnaby, B.C. V5A 1S6

CANADA

Tel: (604) 291-3281

ABSTRACT

Linear modulation methods possess good spectral efficiency. However, their fluctuating envelopes, in conjunction with typically nonlinear power amplifiers, lead to spectral spreading and adjacent channel interference. One remedy is linearization of the amplifier by means of a predistorter. This paper presents and analyzes an adaptive predistorter which is particularly well suited to baseband implementation with DSP hardware. In comparison with the most powerful previously published predistorter, it requires four orders of magnitude less memory, reduces convergence time by over three orders of magnitude, eliminates reconvergence time following a channel switch, and eliminates the need for a phase shifter in the feedback path.

January, 1990

This work was funded by the Province of British Columbia through the Science and Technology Development Fund

I. INTRODUCTION

In the past, both theory and practice of mobile communications have emphasized constant envelope modulation, such as FM or GMSK. These techniques allow power amplifiers to be operated in the nonlinear region near saturation, for power efficiency, yet they do not generate intermodulation products in nearby channels. However, continuing pressure on the limited spectrum available is forcing the development of more spectrally efficient linear modulation methods, such as 16QAM and QPSK with pulse shaping. Since their envelopes fluctuate, these methods generate intermodulation products in a nonlinear power amplifier. In the mobile environment, restrictions on out-of-band emissions are stringent, and the designer is faced with two alternatives: back off an inefficient Class A amplifier to an even more inefficient, but linear, operating region; or linearize the amplifier.

This paper describes a method for linearizing a power amplifier by predistorting its input. It is particularly well suited to baseband implementation using digital signal processing techniques, though hybrid variants are easily concocted. It has a number of advantages compared to previously published techniques, as identified below in this section; but briefly, it is powerful and economical, and it adapts rapidly to amplifier or oscillator changes.

Figure 1 shows a generic model for many adaptive amplifier linearization methods. All signal designations refer either to complex baseband signals or the to the complex envelope of bandpass signals. The linearizer creates a predistorted version $v_d(t)$ of the desired modulation $v_m(t)$, making use of its measurements $v_f(t)$ of the actual amplifier output $v_a(t)$. The quadrature modulator creates a real bandpass signal from the components of $v_d(t)$ for input to the power amplifier (PA). The feedback path directs a portion of the real bandpass PA output to a quadrature demodulator for recovery of the complex envelope. Its output $v_f(t)$ is a scaled, rotated, and possibly delayed version of $v_a(t)$. Note that the same oscillator is used in up and down conversion for coherence, and that some methods require a phase shifter for stability.

Linearization by Cartesian feedback [1,2,3] has the great virtue of simplicity; the amplifier input complex envelope $v_d(t)$ is proportional to the difference between the desired $v_m(t)$ and the measured amplifier output $v_f(t)$, as in a classical control system. However, its linearity and its bandwidth (i.e. the gain-bandwidth product) are critically dependent on loop delay. It is therefore ineffective for travelling wave tube amplifiers, or if there is additional analog or digital filtering at any point in the loop. Moreover, its stability depends on precise adjustment of the phase shifter, an adjustment that depends on the current channel in use.

A more robust alternative to Cartesian feedback is a predistorter (PD), in which the linearizer produces $v_d(t)$ by applying $v_m(t)$ to a memoryless nonlinearity complementary to that of the power amplifier. Feedback is used only for adaptation of the predistorter nonlinearity, rather than for real time calculation of $v_d(t)$. Because it is insensitive to loop delay, predistortion has attracted considerable attention. A few versions are outlined below.

Several predistorters [e.g. 4,5] adaptively cancel the dominant 3rd order component of the PA nonlinearity. The complexity of such structures increases rapidly if extended to 5th or 7th order terms. Moreover, Class AB amplifiers have significant kinks in the transfer characteristic at low levels which are not well modelled by a cube law. [4] is noteworthy, because the quantity fed back is not required to be an accurate replica of the PA output $v_a(t)$, but rather a simple estimate of the power in the 3rd order spectral skirts, as measured by narrowband filters and envelope detection. This, however, restricts its use to modulation of a specific bandwidth.

Another predistorter [6] is related to the one in the present paper, in that it is based on complex gain. However, it requires a dynamic phase shifter, and is not adaptive.

A number of predistorters [e.g. 7,8] are restricted to particular modulation formats. Implemented in digital baseband as RAM lookup tables (LUTs), with an entry for each predistorted point in the signal constellation, they are fast and require very little memory. However, they are limited to rectangular pulses, or to pulse shaping implemented

by filtering following the PA, an unattractive arrangement. [8] is an adaptive version, which requires conversion between polar and rectangular representations.

The most general and powerful predistorter to date was reported by Nagata [9]. It generalizes the table lookup approach of [7,8] to provide the predistorted equivalent v_d of any input value v_m (Figure 2), thereby mapping the complex plane into itself. It is therefore unrestricted by order and type of PA nonlinearity (provided it is memoryless). Since this permits pulse shaping prior to predistortion, it is not restricted by modulation format, either. Nagata also provided an update algorithm for adaptation of the table, and a delay compensation algorithm so $v_m(t)$ and $v_a(t)$ can be compared. We will refer to this technique in the present paper as the "mapping predistorter".

Powerful though it is, the mapping PD has several drawbacks. The lookup table (LUT) is 2 Mword long for 10 bit representation of the real and imaginary parts of v_m, and increases to 8 Mword for 11 bit representation. It requires a phase shifter in the feedback path for stability in the adaptation update, and convergence is very slow (10 sec, at 16 ksym/sec). Moreover, switching to a new channel requires readjustment of the phase shifter, and reconvergence of the table over another 10 sec.

The new PD presented here, like the mapping PD, is unrestricted by modulation format, or by order of PA nonlinearity. In addition, it has some major advantages compared with the mapping PD. It requires over 4 orders of magnitude less table memory (typically under 100 complex word pairs). It reduces convergence time by a similar factor, to a few msec at 16 ksym/sec. It eliminates the reconvergence time following a channel switch, since the table for each channel is so small that it can be downloaded or simply retained in memory. Finally, it eliminates the need for a phase shifter in the feedback loop.

Section II describes the predistorter structure, and demonstrates its ability to suppress intermod products using only a small table. Section III analyzes the effect of PD nonidealities (especially limited table size) on the PA output; to the author's knowledge, such an

analysis is missing from all previously published descriptions of linearizers. Section IV introduces a fast adaptation algorithm. Finally, Section V summarizes the results and their implications.

II THE GAIN BASED PREDISTORTER

We can model the amplifier as a memoryless nonlinearity in several ways. The most productive for our purposes is as a level dependent complex gain. That is, the complex envelope of its input v_d and output v_a are related by:

$$v_a = v_d\, G(|v_d|^2) = v_d\, G(x_d) \tag{1}$$

where x_d denotes the squared magnitude of v_d, and $G(x_d)$, the complex gain of the amplifier, summarizes its AM/AM and AM/PM characteristics. Note that $G(x_d)$ depends only on the instantaneous power of the input, not on its phase. Figures 3 and 4 show the relation between bench measurements and the complex gain for a typical Class AB amplifier. The effect of compression is clearly evident at high input levels, as is the loss of gain at low levels, because of the crossover point between the push and pull halves of the amplifier.

The gain based PD is not just modelled, but actually constructed, according the the complex gain formulation (Figure 5). Its input and output complex envelopes are related by:

$$v_d = v_m\, F(|v_m|^2) = v_m\, F(x_m) \tag{2}$$

where x_m denotes the squared magnitude of v_m.

For any input power, the optimum value of the PD complex gain F is determined by equating the composite PD/PA nonlinearity to a nominal constant amplitude gain K. Normally, K is selected to be a little less than the amplifier's midrange gain. However, as shown in Section III, the choice of K has no effect on the signal to quantization noise ratio at the amplifier output. Combining (1) and (2), we define F implicitly by:

$$v_m\, F(|v_m|^2)\; G\left(|v_m|^2 |F(|v_m|^2)|^2\right) = Kv_m \tag{3}$$

or:

$$F(x_m)\; G(x_m |F(x_m)|^2) = K \tag{4}$$

A fast technique for adaptive calculation of $F(x_m)$ is described in Section IV. Note that the PD complex gain $F(x_m)$ has a real domain, so that it can be represented by a one-dimensional LUT, rather than the

two-dimensional LUT required by the mapping PD.

In practice, there is little point in trying to linearize an amplifier up to its saturated output power P_{sat}, because the distortion increases drastically in this region, and we are faced with rapidly diminishing returns on linearization effort. Accordingly, we define the span S as the fraction of saturated power over which we attempt linearization. Thus the maximum output power is given by $S\,P_{sat}$. Realistic values for the span are in the range 0.95 to 0.98. An alternative description is the peak backoff (PBO) of the PA in dB:

$$PBO = -10 \log(S) \tag{5}$$

The span in turn limits the domain of the linearizer:

$$0 \le x_m \le P_{mm} \tag{6}$$

where the maximum power of v_m is given by:

$$P_{mm} = S\,P_{sat}/K^2 \tag{7}$$

This paper examines the implementation of $F(x_m)$ as a LUT with entries equispaced in input power x_m. Although a good case can be made for nonuniform spacing, as we shall see in Section III, it would be at the cost of more computation when implemented digitally. With N_t table entries, the step size is given by:

$$D_g = P_{mm}/N_t \tag{8}$$

The range and midpoint of each step, and the corresponding table entries, are given for $i = 0, 1, \ldots, N_t-1$ as:

$$X_{mi} = \{x_m: iD_g \le x_m < (i+1)D_g\} \tag{9}$$

$$x_{mi} = D_g(i + \tfrac{1}{2}) \tag{10}$$

$$F_i = F(x_{mi}) \tag{11}$$

That is, the table is optimized according to (4) for the midpoint of each step, as shown in Figure 6.

The remainder of this section demonstrates that even very small PD gain tables can produce major reductions in intermodulation products,

using in all cases the Class AB amplifier of Figures 3 and 4 with a PBO of about 0.22 dB, giving a 95% span. First we define the output backoff (OBO) as the dB difference between P_{sat} and the average signal output power. Clearly it depends on the peak to average power ratio of the signal, which we call the signal backoff (SBO):

$$OBO = PBO + SBO \tag{12}$$

Large OBO values imply inefficient operation, due either to amplifier peak backoff or to the signal format. Note that compression lowers the SBO below the value characteristic of the signal in a linear regime.

Our first example is a two-tone test. Although it does not fully exercise the nonlinearities, it is widely used. Its SBO is 3 dB in the absence of distortion. Figure 7 shows the results of two-tone tests of the amplifier with PBO values of 0.22 dB and 24.63 dB (i.e. SBO values of 2.53 and 27.67 dB, respectively). All spectra are normalized to 0 dB for the desired components at $f/f_s = \pm 0.0156$, and the frequency is normalized by the sampling rate. Even with a PBO of 25 dB, the dominant IM products of the uncorrected amplifier are only 40 dB down. In contrast, the third curve on the same figure shows that a 64 point predistortion table achieves the desired PBO of 0.22 dB (OBO of 3.22 dB) with IM products reduced below 60 dB.

The second example (Figure 8) is a more demanding noise loading test. Thirty sine waves with equal amplitudes and randomly selected phases are located fifteen on each side of a center channel spectral null. The degree to which the null fills in at the output of the PA indicates the total power in the nonlinear products. Averaging of a few such output spectra, each with a new set of phases, is sufficient. The SBO for this signal is 14.8 dB. With a SBO of 13.2 dB (PBO of 0.22 dB), wide IM skirts about 14 dB down are evident in the uncorrected amplifier, and even an OBO of 33.7 dB lowers the IM by only about 8 dB. In contrast, both a 32 point and a 64 point PD table introduce a major reduction in intermod power at an OBO of 15 dB (PBO of 0.22 dB).

The final example is more closely related to data transmission. The input signal is 16QAM, with a square root spectral raised cosine pulse with 25% rolloff, Hamming windowed to 7 symbols. The SBO is about 6 dB. As shown on Figure 9, the 3rd, 5th and 7th order skirts are

clearly visible when the uncorrected amplifier is operated at 0.22 dB PBO (3.0 dB OBO). Even with an OBO of 33.7 dB (PBO of 30.24 dB), out of band emission is still about 56 dB down, and the power efficiency of the amplifier is hopelessly small. Again we see a striking improvement with the gain based PD operated at a PBO of 0.22 dB: even a 32 point gain table brings the skirts below 60 dB, and a 64 point table lowers the IM by another 6 dB (see Section III).

These examples demonstrate that major improvements in linearity can be achieved with very small tables: N_t = 64 entries, compared with over 1 million entries in the mapping PD. This dramatic reduction is due to exploitation of the rotational invariance of the amplifier nonlinearity. Of the two dimensions required to specify a point in the plane, only one (radius) need be quantized; the other (phase) remains continuous.

The cost, compared with the mapping PD, is computation. Both the squared magnitude of v_m (for the table index) and the complex multiplication of v_m by the table entry have to be performed at a sampling rate adequate to represent the highest order nonlinearity of interest. As a point of interest, we have implemented such a PD on a TMS320C25 at 240 kHz, which is sufficient for 7th order nonlinearities when the nominal channel bandwidth is 30 kHz.

III ERROR ANALYSIS OF THE PREDISTORTERS

The previous section presented experimental evidence that the gain based PD can achieve remarkable reduction in IM products with a table four orders of magnitude smaller than that of the mapping PD. In this section we provide an analytical explanation for the results, by exploring the relation between PA characteristics and table size. Among other findings, we demonstrate: that table size effects show up as relative error in the gain PD, compared with absolute error in the mapping PD; that error power in the gain PD decreases inversely as N_t^2, rather than inversely as N_t in the mapping PD; and that the jitter in the gain table entries caused by adaptation can be compensated by a relatively small increase in table size.

A. The Mapping Predistorter

In the mapping PD, the input v_m, representing the desired output of the amplifier, is quantized to several bits in its real and imaginary parts separately, and acts as an index to the table. Tne table size is determined by the number of bits; for example, with 10 bit accuracy, 2^{20} or about 1 million table entries are needed. Note that the number of bits in each table entry is a different issue. For simplicity, we will assume infinite precision in the table.

Analysis of the effects of input quantization through two nonlinearities is surprisingly straightforward. Imagine the input space (v_m space) as partitioned into quantization cells, as shown in Figure 10. If there are N_t entries in the two-dimensional table, then each cell has width:

$$D_m = \sqrt{2\, P_{mm} / N_t} \qquad (13)$$

since we associate maximum power P_{mm} with the extreme corner cells. All v_m values in a cell are mapped onto a single v_d value, which we assume to be the correct predistorter value for the center of the cell. Thus the amplifier output v_a has its desired value Kv_m only at the cell centers. At other points in the cell there is a quantization error e_m (where the subscript denotes the mapping PD):

$$v_a = Kv_m + e_m \tag{14}$$

which is bounded in its real and imaginary parts by $KD_m/2$:

$$|\mathrm{Re}[e_m]| \le KD_m/2 \qquad |\mathrm{Im}[e_m]| \le KD_m/2 \tag{15}$$

Now we make the usual assumptions in quantization analysis: the error e_m is a random variable independent of v_m, uniformly distributed over a cell, and its real and imaginary parts are independent. With these assumptions, we can calculate its variance:

$$\sigma_{em}^2 = \frac{1}{2} E[|e_m|^2] = K^2 D_m^2/12 \tag{16}$$

or, in terms of table size:

$$\sigma_{em}^2 = K^2 P_{mm}/6N_t \tag{17}$$

We can calculate a SNR if we know the statistics of v_m. Assume for simplicity that v_m is also uniformly distributed over its domain, in which real and imaginary parts are bounded in magnitude by $\sqrt{P_{mm}/2}$. Then SNR_m, the ratio of signal to quantization noise powers at the PA output with a mapping PD, is given by:

$$SNR_m = K^2 \sigma_{vm}^2/\sigma_{em}^2 = N_t \tag{18}$$

A more common measure of out-of-band emission is the ratio of peak power spectral densities (PSDs) of signal and noise. Since the additive error appears as white noise, we can estimate this quantity as:

$$PSD_{va}/PSD_{em} \cong SNR_m \, W_m/f_s \tag{19}$$

where W_m is the two-sided bandwidth of the modulation, and f_s is the sampling rate at which the predistorter operates.

In summary, the effect of finite table size in the mapping PD is to add independent white noise, for which the SNR is proportional to the table size N_t.

B. The Gain Based Predistorter

The gain based predistorter described in Section II has a table indexed by input power x_m. Thus the quantization cells in v_m space are annular regions, or rings. Because the table entries are equispaced in power, the rings are more closely spaced in amplitude as the input level increases. This improves representation accuracy in the region near saturation, where amplifier distortion usually varies most rapidly. Unfortunately, it does little for the distortion at low input levels characteristic of Class AB amplifiers.

Before examining the effect of table size, we establish a useful relation. Since (4) determines the correct value of F at every x_m, we also have:

$$\frac{d}{dx_m} F(x_m)\ G(x_m|F(x_m)|^2) \equiv 0 \tag{20}$$

Carrying through the differentiation gives:

$$F'(x_m)\ G\left(x_m|F(x_m)|^2\right) + F(x_m)\ G'\left(x_m|F(x_m)|^2\right)\ \left(|F(x_m)|^2 + 2\ x_m\ \mathrm{Re}\left[F'(x_m)\ F^*(x_m)\right]\right) \equiv 0 \tag{21}$$

where the dash indicates a derivative and the asterisk a conjugate.

Now to determine the effect of quantization of F(x), that is, the effect of a finite number N_t of table entries. At any point, imagine that the ideal value of F is perturbed by a small error ΔF. This changes the input to the PA and therefore induces a small change ΔG in G. The PA output is now given by:

$$\begin{aligned} v_a &= v_m\ (F + \Delta F)\ (G + \Delta G) \\ &\cong v_m\ (K + G\Delta F + F\Delta G) \\ &= v_m\ \left(K + G\Delta F + 2xFG'\mathrm{Re}\left[F^*\Delta F\right]\right) \end{aligned} \tag{22}$$

where F, G, F′ and G′ denote quantities evaluated before the perturbation. Note that the error in F induces a relative error in v_a, in contrast to the absolute error we saw earlier in the mapping PD.

Next, we relate the error in F to the table size. Recall from Section II that the table entry is optimized for the midpoint of the step. The error in F is approximately proportional to Δx_m, the

difference between actual input power and the midpoint of the step:

$$\Delta F \cong F'(x_m)\,\Delta x_m \tag{23}$$

Substitution of (23) into (22) gives the perturbed amplifier output:

$$v_a \cong v_m \left[K + \left(GF' + 2xFG'\mathrm{Re}\left[F^* F'\right]\right)\Delta x_m\right] \tag{24}$$

which, from (21), can be rewritten as:

$$v_a \cong v_m \left[K + F|F|^2 G'\,\Delta x_m\right] = Kv_m + e_g \tag{25}$$

where e_g is the quantization error at the PA output with a gain based PD. Substitution of the basic inverse relationship (4) gives expressions for the output quantization error e_{gi} in the i^{th} step in terms of PA characteristics only:

$$e_{gi} \cong v_m F_i |F_i|^2 G_i' \Delta x_m = v_m K^3 \frac{G_i'}{G_i |G_i|^2} \Delta x_m \tag{26}$$

where G_i and G_i' stand for $G(x)$ and $G'(x)$, respectively, evaluated at:

$$x_{di} = x_{mi} |F_i|^2 \tag{27}$$

To determine the SNR at the amplifier output, assume that there are enough table entries to make Δx_m a random variable uncorrelated with x and uniformly distributed over $[-D_g/2, D_g/2)$. Then the variance of e_{gi} evaluated over the Δx_m ensemble is:

$$\sigma_{egi}^2 = \tfrac{1}{2} E[|e_{gi}|^2]$$

$$= \tfrac{1}{2} x_{mi} K^6 \frac{|G_i'|^2}{|G_i|^6} \frac{D_g^2}{12} \tag{28}$$

Substitution of (7) and (8) for D_g gives:

$$\sigma_{egi}^2 = \frac{1}{24} \left(\frac{S\, P_{sat}}{N_t} \right)^2 \frac{|G_i'|^2}{|G_i|^6} x_{ai} \tag{29}$$

where $x_{ai} = K^2 x_{mi}$. Note that the error variance expressed this way in terms of amplifier output quantities does not depend on the nominal

amplitude gain K. The average quantization noise power can be written:

$$\sigma_{eg}^2 = \sum_i p_i \, \sigma_{egi}^2 \tag{30}$$

where p_i is the probability density function of the levels x_{ai} induced by the particular modulation, such as QPSK or 16QAM. The overall SNR in the gain based PD is therefore given by:

$$SNR_g = \sigma_{va}^2 / \sigma_{eg}^2 = \frac{1}{2} \sum_i p_i \, x_{ai} \Big/ \sum_i p_i \, \sigma_{egi}^2 \tag{31}$$

As in the mapping case, the ratio of spectral densities of signal and additive quantization noise depends on bandwidth and sampling rate:

$$PSD_{va}/PSD_{eg} \cong SNR_g \, W_m/f_s \tag{32}$$

The variance in the output quantization error is inversely proportional to the square of the table size, as shown by (29) and (30). This contrasts with the mapping based PD, in which error power varies as the simple inverse of table size. The variance also depends on the AM/AM and AM/PM characteristics of the amplifier through G_i and G_i', and on the modulation, through p_i. Again this differs from the mapping PD, in which the SNR depends on neither of these characteristics.

Figure 11 illustrates the variation of σ_{egi}^2 as a function of level for a 32-point gain table and the Class AB amplifier of Figures 3 and 4. The quantization error is negligible over most of the input range. As expected, the kink at very low input levels contributes the largest error. Thus modulation techniques which avoid very low levels, such as QPSK with $\pi/4$ precession, can substantially reduce the IM spectral skirts. However, SNR_g is clearly in the acceptable range, even if the low levels are included. Also shown is the relative error variance, obtained by dividing σ_{egi}^2 by x_{ai}. It, too, rises sharply at low levels.

Finally, note that there are ways to reduce the error further. First, doubling the table size reduces it by a factor of 4, or 6 dB. Second, nonuniform steps in the gain table would allow us to represent both low and high amplitudes more accurately; however, this is at the cost of additional computation, since determining the correct table index from x is more difficult. Finally, from (32), we can decrease the

IM PSD relative to the signal by increasing the sampling rate f_s. Again, this is at the cost of additional computation.

C. Effect of Adaptation Jitter

We demonstrated above that output quantization noise is negligible in the case of a fixed predistorter, optimized at the center point of each quantization step. An adaptive PD, however, is not optimized for that value. As we will see in Section IV, the successive Δx_m values are scattered over the range $[-D_g/2, D_g/2)$, and the adaptation algorithm chases a moving target. This causes the table entry to jitter, and the output error to increase. This section demonstrates that small increases in table size are sufficient to compensate for adaptation jitter.

First, consider the worst case: successive Δx_m values alternate between the two extremes of $-D_g/2$ and $D_g/2$, and the algorithm converges so quickly that the table entry is always optimized for the previous Δx_m value. The situation is similar to the fixed predistorter considered above, with Δx_m always taking the values $\pm D_g$. The variance is therefore D_g^2, rather than $D_g^2/12$, as in the earlier case. The twelve-fold increase in error variance can be compensated by increasing the number of entries in the table by a factor of $\sqrt{12}$, about 3.5.

A more reasonable assumption is that successive Δx_m values are independent, and uniformly distributed over $[-D_g/2, D_g/2)$. Pessimistically again, we assume that the convergence algorithm leaves the table entry optimized for the previous value. This situation is similar to the fixed predistorter with Δx_m equal to the difference of two uniformly distributed variates; that is, its pdf is triangular over $[-Dg, Dg)$:

$$p_{\Delta xm}(\xi) = \begin{cases} \left(1 - |\xi|/D_g\right)/D_g, & |\xi| \le D_g \\ 0, \text{ elsewhere} \end{cases} \qquad (33)$$

Its variance is therefore $D_g^2/6$, that is, twice the variance of the fixed predistorter case. The output quantization error power doubles, but this can be compensated simply by increasing the number of table entries by a factor of $\sqrt{2}$.

A still more accurate model would account for the averaging effect

introduced by a realistic adaptation algorithm, one without instant convergence. Since Δx_m is uniformly distributed, F_i will tend to be close to its optimized midpoint value, and the required increase in table size will be relatively small, on the order of about 20%.

IV. ADAPTATION OF THE PREDISTORTER

General Comments

Amplifier characteristics drift because of aging, temperature changes, supply voltage variations, and channel switches. Generally, this means that if spectral emission constraints are stringent enough to require the use of a predistorter, then the PD will have to be adaptive.

Several previously reported adaptive predistorters are cited in Section I. In all cases, though, they display linear convergence; that is, they add roughly the same number of bits of accuracy in each iteration. In this section, we demonstrate that much faster convergence can be obtained with algorithms made possible by the power of a digital signal processing (DSP) implementation. In addition, we can discard the phase shifter or PLL in the feedback path.

There are several issues of importance in adaptation. For example, three scenarios of interest are initial convergence with an unknown amplifier, reconvergence after a channel switch, and tracking slow drift. Performance issues include speed, stability and computational load.

A. Linear Convergence Techniques

In addition to the mapping PD itself, Nagata [9] reported an adaptation algorithm. Its combination of reasonable speed and reasonable computational load appears to make it the best so far of the linear convergence methods.

Each table entry is updated when the input v_m passes through the corresponding quantization cell. Figure 12 shows that the iteration update for a given cell is structured as a feedforward path with a correction signal v_c. Note that the time step in the delay block corresponds to an iteration, not a sample time. Apart from the nonlinearity of the power amplifier, the loop resembles a classical feedback control system, and it can be analyzed in the same way. When it converges, therefore, the error decreases exponentially with iteration number; that is, it adds roughly the same number of bits of accuracy with each iteration.

We experimented with this algorithm using the Class AB amplifier introduced earlier. The convergence of the squared magnitude of the output error e_m for selected parameters is shown in Figure 13. The iteration parameter a was set at 0.4; larger values resulted in faster convergence at low amplitudes, but failure at high amplitudes. The curves demonstrate that convergence is slow at high output power levels (30 watt is about 0.22 dB PBO). Moreover, the phase shift ϕ must be within about 0.5 rad of its optimum value -1.7 rad, or the iteration diverges. With an additional 3 dB backoff (15 watt output), convergence is faster, and less sensitive to phase shift (though of course ϕ must remain within about 1.5 rad of its optimum value to avoid positive feedback).

The phase shift can be implemented either in DSP at baseband, or as a PLL in the down conversion chain, as indicated in [9]. In either case it represents an additional load of computation or hardware.

The large number of table entries (about 10^6 complex values) results in very slow convergence of the PD as a whole, since each entry exercised by the modulation scheme must be iterated many times, as demonstrated above. The estimate in [9] is 10 sec for binary modulation at 16 ksym/sec, and our simulation suggests this is optimistic. To make matters worse, a switch of channel normally results in a phase jump in the loop, which forces reconvergence, with another 10 sec delay.

In favour of this adaptation algorithm is the low computational load: each iteration requires 3 complex additions, a real by complex multiplication and no divisions.

B. Secant Method

To update the table entries in the gain based PD, we formulate the adaptation as a root finding problem to which a number of classical methods can be applied. From (25), the desired value of F for a given v_m is the root of:

$$e_g(F) = v_a(F) - Kv_m \tag{34}$$

a quantity which can be measured as the difference between the actual and the desired amplifier output. Ideally, each table entry F_i is a root of (3) if the squared magnitude of v_m equals x_{mi}, the midpoint of

the quantization step.

Since the derivative of $e_g(F)$ is not available, we cannot apply Newton's method [10] to (34) and obtain quadratic convergence. A reasonable alternative, which offers convergence speed intermediate between linear and quadratic convergence, is the secant method [10]:

$$F_i(k+1) = \frac{F_i(k-1)e_g(F_i(k)) - F_i(k)e_g(F_i(k-1))}{e_g(F_i(k)) - e_g(F_i(k-1))} \qquad (35)$$

where $F_i(k)$ is the k^{th} iteration of table entry i. The iteration update is applied whenever v_m falls into the i^{th} quantization step. It is easy to show [10] that the number of bits of accuracy in F_i increases by a factor of 1.6 at each iteration in the neighbourhood of the root, which is considerably faster than the linear convergence we examined earlier.

The iteration (35) is familiar for real roots of real functions. Strictly speaking, it is the true secant method in the complex domain only if $e_g(F)$ is analytic. However, from (3) we can write:

$$e_g(F) = v_m F\, G\left(x_m |F|^2\right) - Kv_m \qquad (36)$$

from which it is evident that $e_g(F)$ is not analytic. We take a pragmatic view at this point: (35) defines an iteration which we found experimentally to work very well.

The convergence of squared magnitude of e_g when the method is applied to the same Class AB amplifier we used earlier is shown in Figure 13. The two starting points for the iteration were $F_i(0) = 0$, and $F_i(1) = .001$; although they are both real, the iteration quickly moves out into the complex plane. Convergence at 30 watts is significantly slower than at 15 watts. This is a consequence of saturation in the amplifier; near the root the derivative (or rather the norm of the Jacobian matrix) also becomes very small, which is a classical difficulty for both Newton's method and the secant method.

The secant method nevertheless converges more quickly than the linear method, taking half the number of iterations to reduce the output error power to 0.1 mwatt. Of more importance, though, is the fact that the same behaviour is obtained for every value of phase shift ϕ, so that the PLL or phase shifter required by the linear method is made

unnecessary.

We can now estimate the time for initial convergence of the predistorter. Since each table entry needs about 10 iterations for convergence, the complete PD takes about 10 N_t iterations. Assume that there is an iteration for each sample of the input. The required sampling rate for a signal of desired RF bandwidth W, and accurate representation of n^{th} order intermod products, is nW (recall that these are complex samples). The time to converge is therefore approximately:

$$T_c = \frac{10\ N_t}{nW} \tag{37}$$

In particular, for 25 kHz channels and 7^{th} order products, and for a table of 64 entries, the convergence time is under 4 msec. Reconvergence time after a channel switch is even shorter, since the table is so small that values from the last use of the new channel can be loaded, and the PD resumes in a virtually converged state. These results represent a major improvement in speed compared with the 10 sec taken by the mapping PD with linear convergence.

The drawback of the method is its computational load. From (35), each iteration requires 2 complex additions, 4 complex multiplications, and 2 complex by real divisions. This is significantly more than in the linear convergence method. However, the load can be distributed by performing updates on only a decimated subsequence of input samples, with a corresponding increase in convergence time. A fertile area for investigation should be the development of iteration algorithms which require less computation, but retain speed and independence of phase shift.

V. CONCLUSIONS

Like the mapping PD, the gain based PD presented here is applicable to any modulation format, and any AM/AM and AM/PM characteristic. However, it has major advantages compared with the mapping PD: it reduces memory requirements from 1 million complex word pairs to under 100 complex word pairs; it reduces convergence time with a new amplifier from 10 sec to about 4 msec; it reduces reconvergence time following a channel switch from 10 sec to virtually zero; and it removes the need for a phase shifter or PLL in the feedback path.

On the other hand, the improvement is at the expense of increased computation, both in the real time multiplication of each sample by the complex gain, and in the adaptation update step.

Other significant findings from the analysis were as follows. Error in the PD table entries appears as a relative error in the gain based PD, compared with an absolute error in the mapping PD. The quantization error variance decreases as the inverse square of table size in the gain based PD, compared with a simple inverse in the mapping PD. Quantization error in the gain based PD, unlike the mapping PD, depends on the amplifier characteristic and on the instantaneous power level at the output; most error is contributed at levels near saturation, and near zero in the case of a Class AB amplifier.

The predistorter presented here is well suited to DSP implementations of the next generation of spectrally efficient mobile communications equipment. Since it allows linearization of the power amplifier up to at least 95% of the saturated output power, it provides power efficiency as well as spectral efficiency.

REFERENCES

[1] Y. Akaiwa and Y. Nagata, "Highly Efficient Digital Mobile Communications with a Linear Modulation Method", IEEE J Sel Areas in Comms, vol SAC-5, no 5, pp 890-895, June 1987.

[2] A. Bateman, D.M. Haines, and R.J. Wilkinson, "Linear Transceiver Architectures", Proc IEEE Vehic Tech Conf, pp 478-484, Philadelphia, 1988.

[3] S. Ono, N. Kondoh and Y. Shimazaki, "Digital Cellular System With Linear Modulation", Proc IEEE Vehic Tech Conf, pp 44-49, San Francisco, 1989.

[4] M. Nannicini, P. Magni, and F. Oggioni, "Temperature Controlled Predistortion Circuits for 64 QAM Microwave power Amplifiers", IEEE MTT-S Digest, pp 99-102, 1985.

[5] J. Namiki, "An Automatically Controlled Predistorter for Multilevel Quadrature Amplitude Modulation", IEEE Trans Commun, vol COM-31 no 5, pp 707-712, May 1983.

[6] H. Girard and K. Feher, "A New Baseband Linearizer for More Efficient Utilization of Earth Station Amplifiers Used for QPSK Transmission", IEEE J Selected Areas in Commun, vol SAC-1, no 1, pp 46-56, January 1983.

[7] J. Graboski and R.C. Davis, "An Experimental MQAM Modem Using Amplifier Linearization and Baseband Equalization Techniques", Proc Natl Commun Conf, pp E3.2.1-E3.2.6, 1982.

[8] A.A.M. Saleh and J. Salz, "Adaptive Linearization of Power Amplifiers in Digital Radio Systems", Bell Syst Tech J, vol 62, no 4, pp 1019-1033, April 1983.

[9] Y. Nagata, "Linear Amplification Technique for Digital Mobile Communications", Proc IEEE Vehic Tech Conf, pp 159-164, San Francisco, 1989.

[10] Germund Dahlquist and Ake Bjorck, Numerical Methods, Prentice-Hall, 1974.

Figure 1: Generic Configuration for Adaptive Linearization
Figure 2: The Mapping Predistorter
Figure 3: AM/AM and AM/PM Characteristics of a Class AB Amplifier
Figure 4: Complex Gain Characteristic of a Class AB Amplifier
Figure 5: The Gain Based Predistorter
Figure 6: Quantization in the Complex Gain Table
Figure 7: Two Tone Test, Showing Effect of Predistorter
Figure 8: Noise Loading Test, Showing Effect of Predistorter
Figure 9: 16QAM Data Signal, Showing Effect of Predistorter
Figure 10: Input Quantization in the Mapping Predistorter
Figure 11: Distribution of Intermod Power With Instantaneous Output Level Using Gain Based Predistorter
Figure 12: A Linear Adaptation Update Structure
Figure 13: Convergence Behaviour of Linear and Secant Update Algorithms

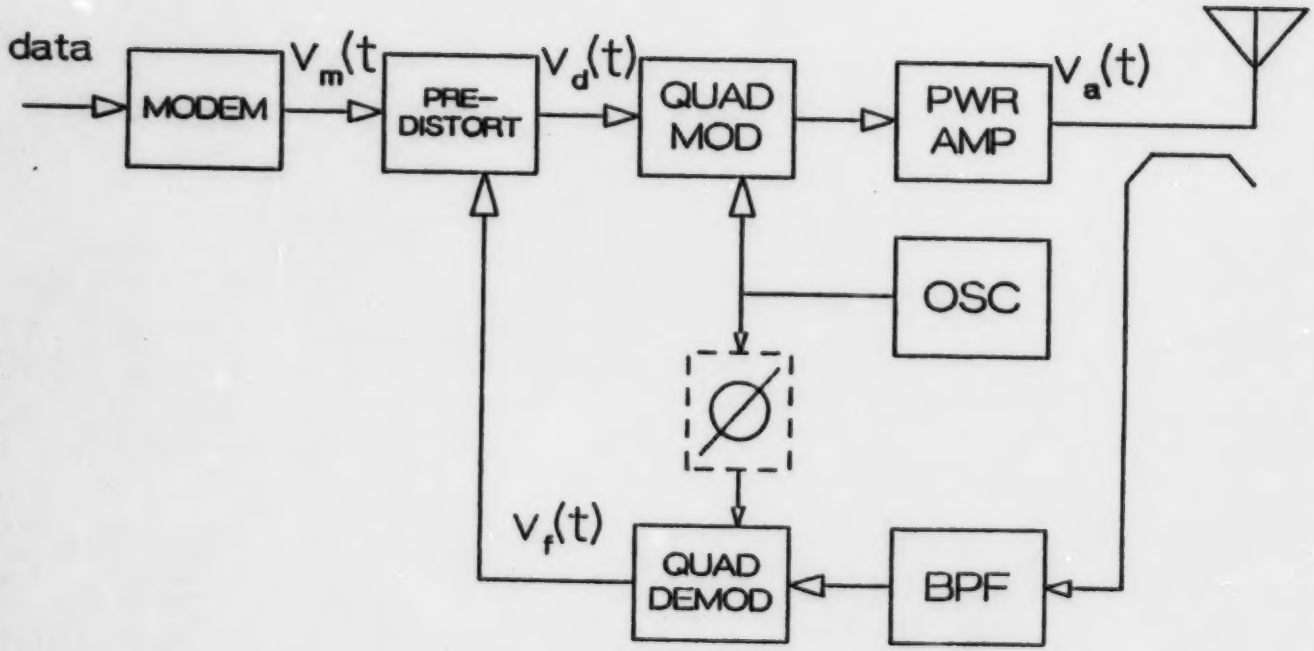

Figure 1: Generic Configuration for Adaptive Linearization

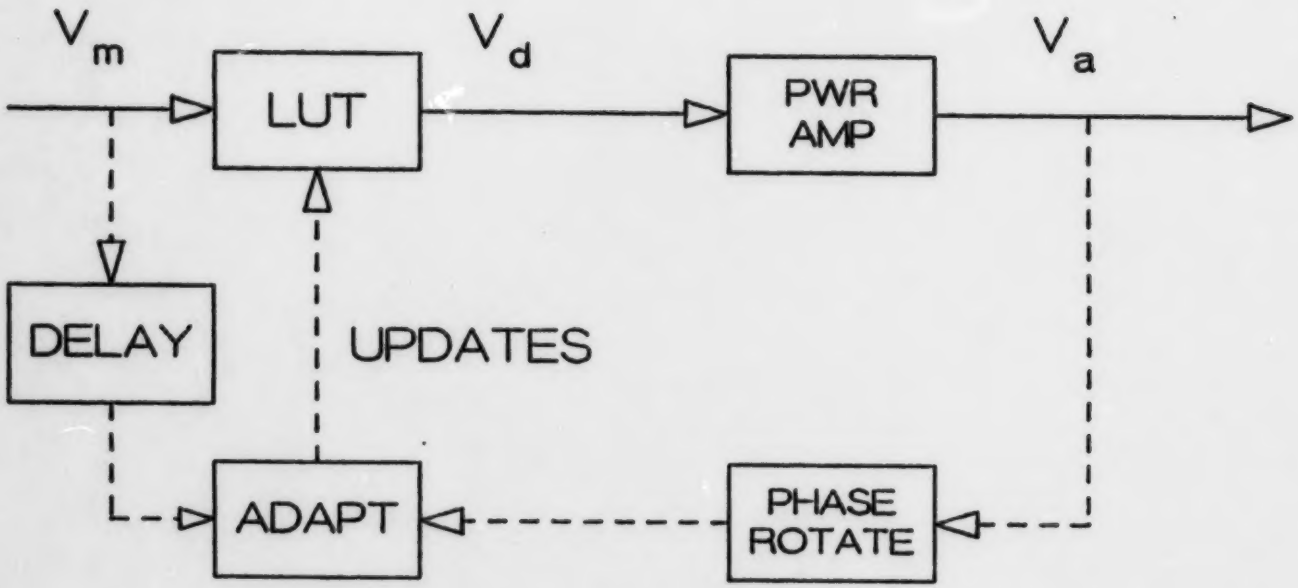

Figure 2: The Mapping Predistorter

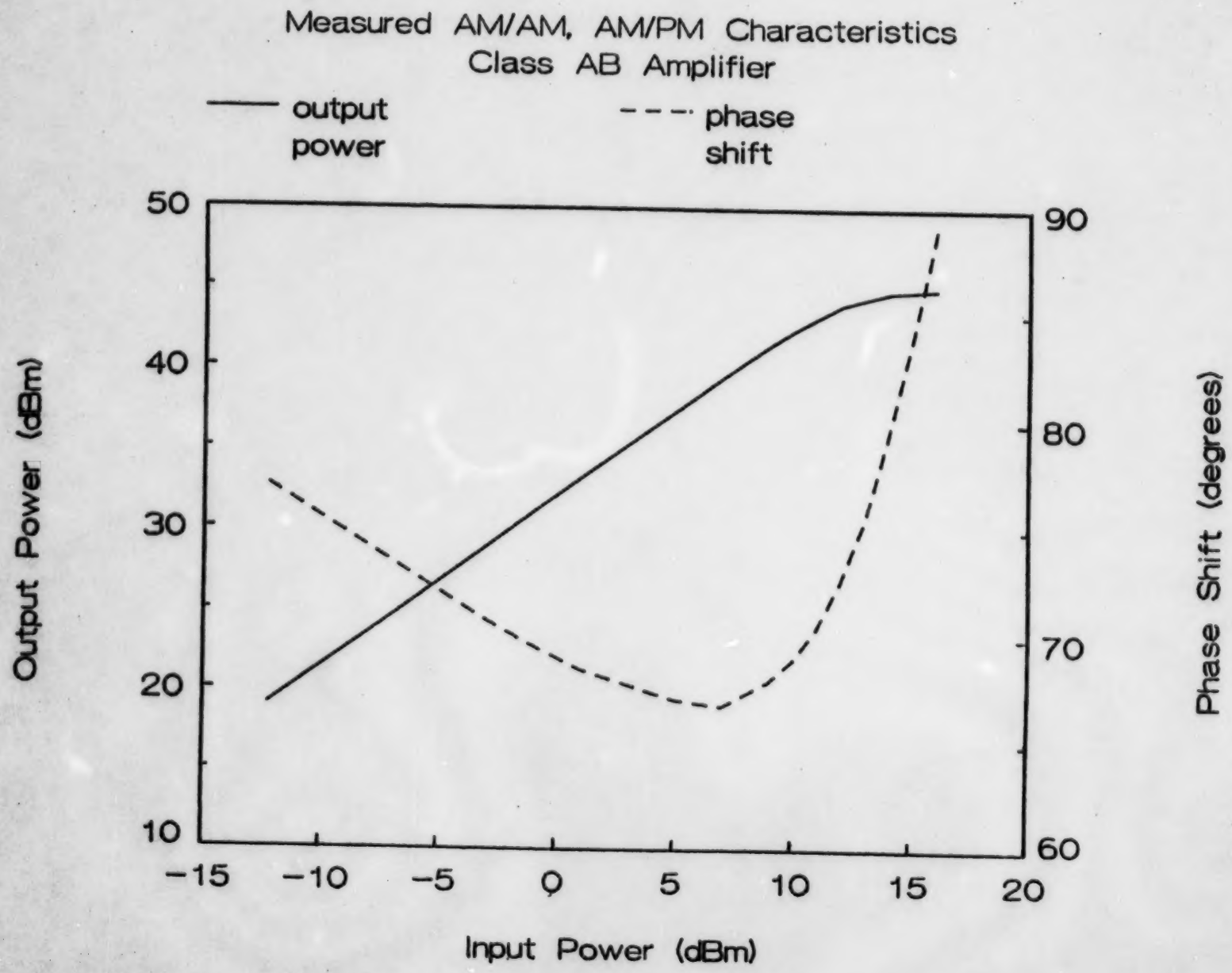

Figure 3: AM/AM and AM/PM Characteristics of a Class AB Amplifier

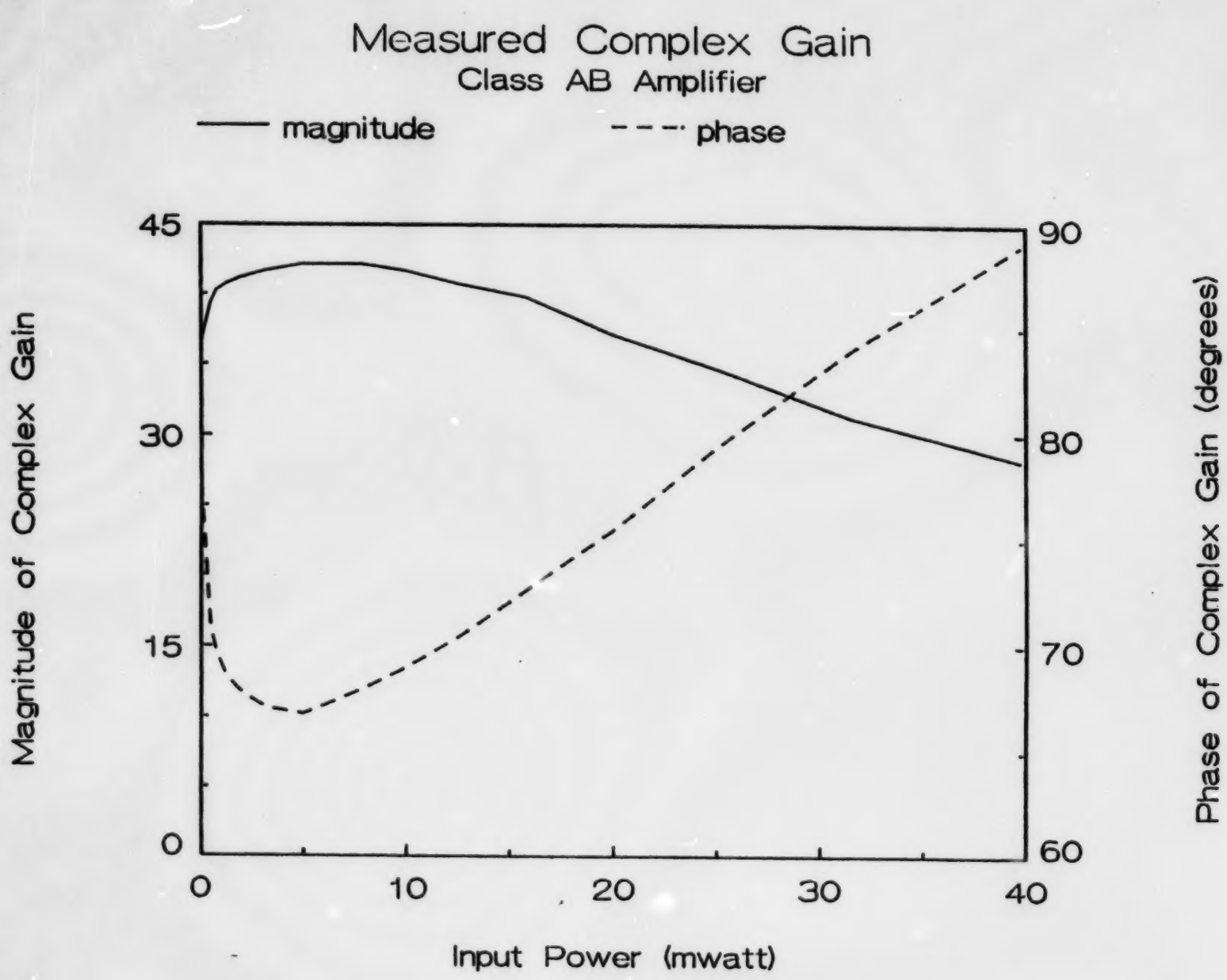

Figure 4: Complex Gain Characteristic of a Class AB Amplifier

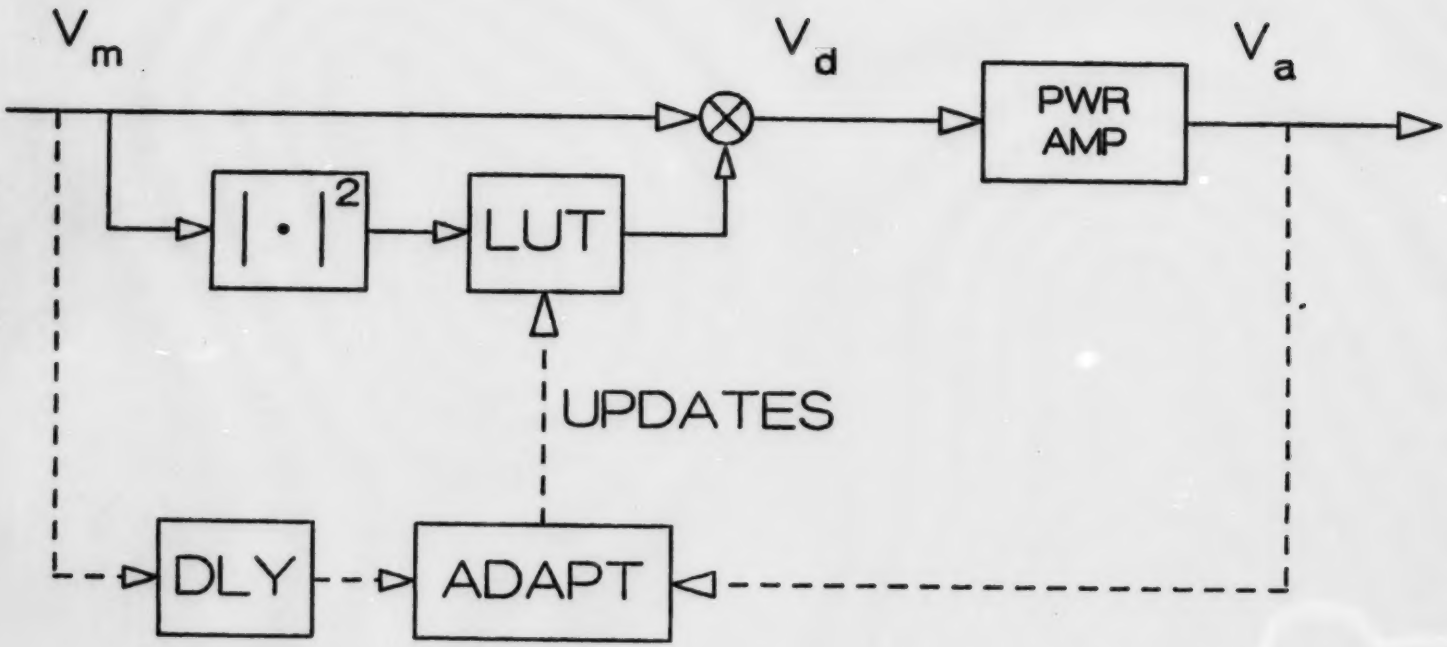

Figure 5: The Gain Based Predistorter

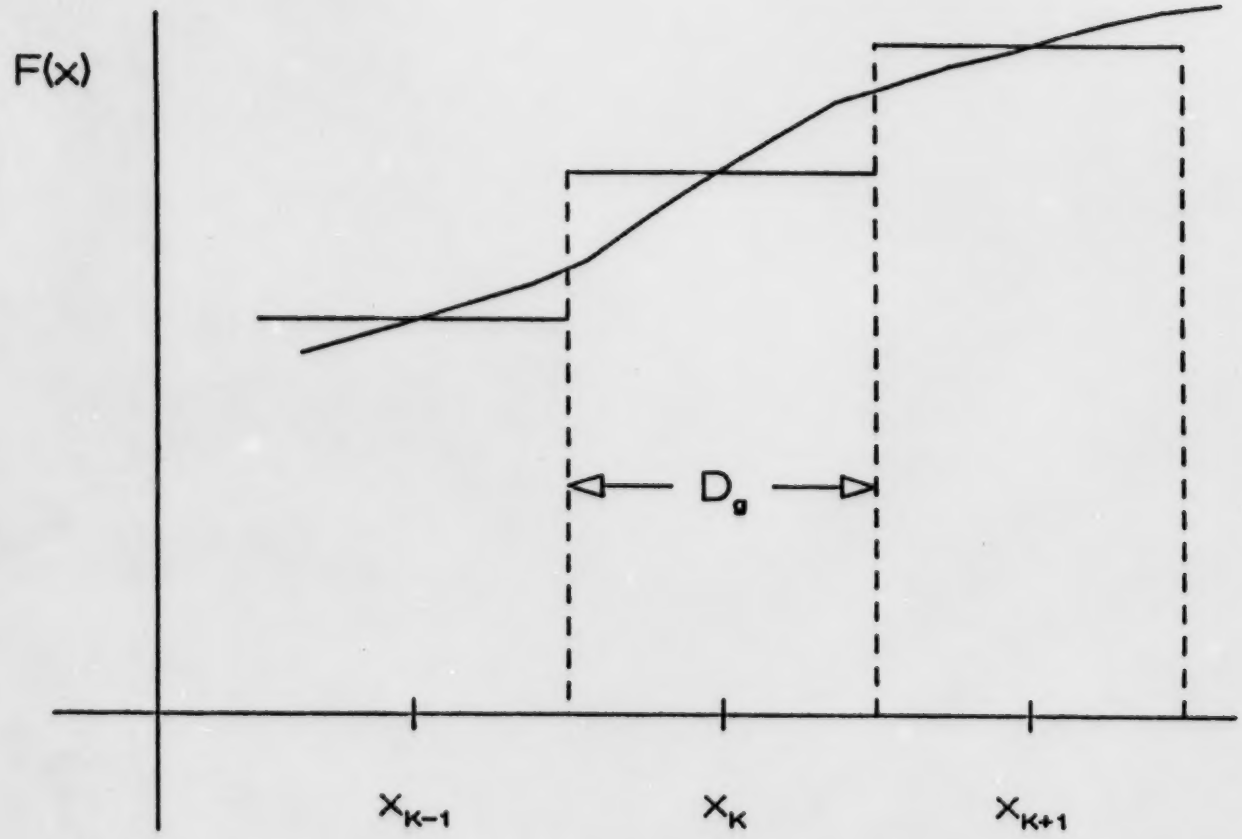

Figure 6: Quantization in the Complex Gain Table

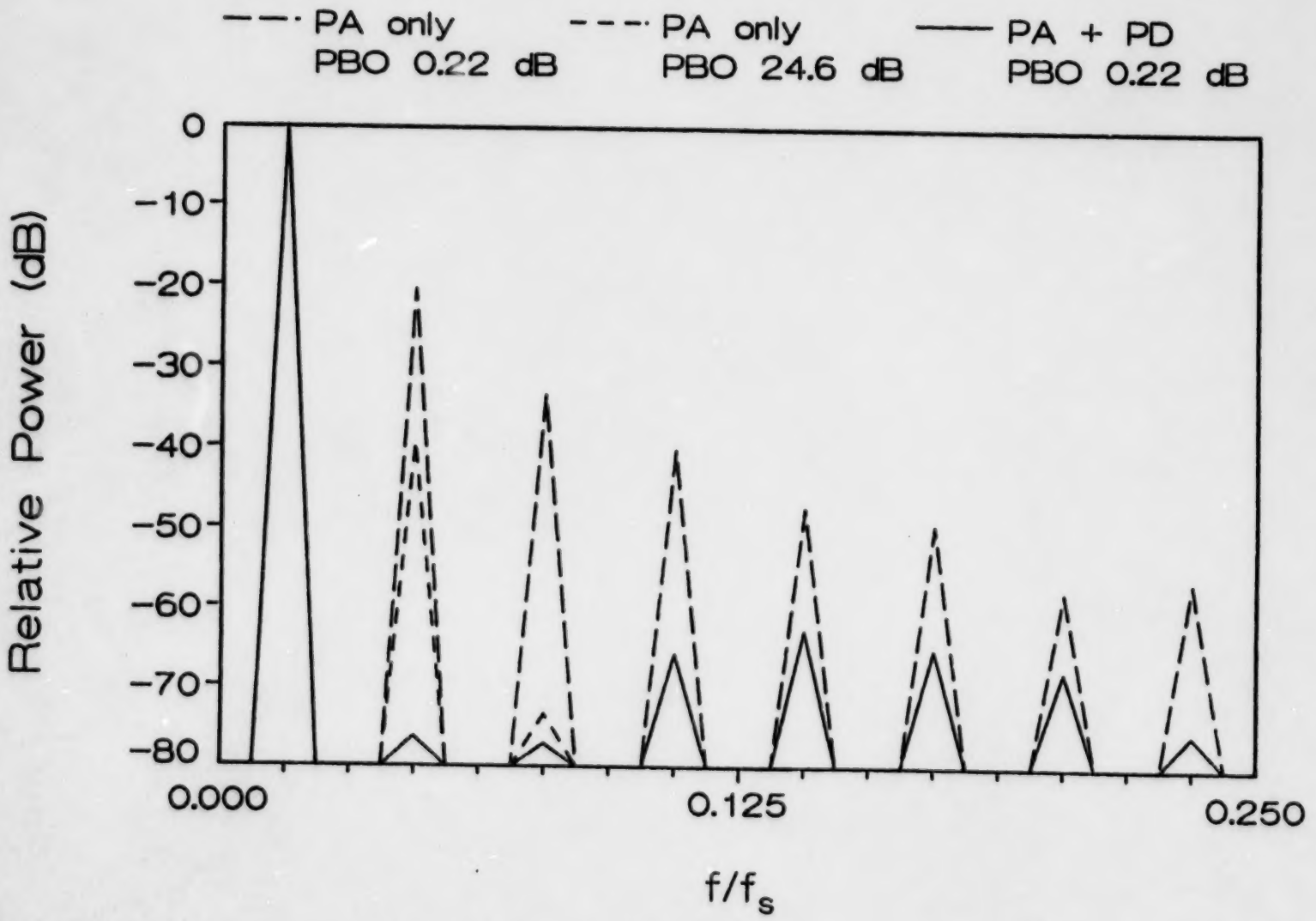

Figure 7: Two Tone Test, Showing Effect of Predistorter

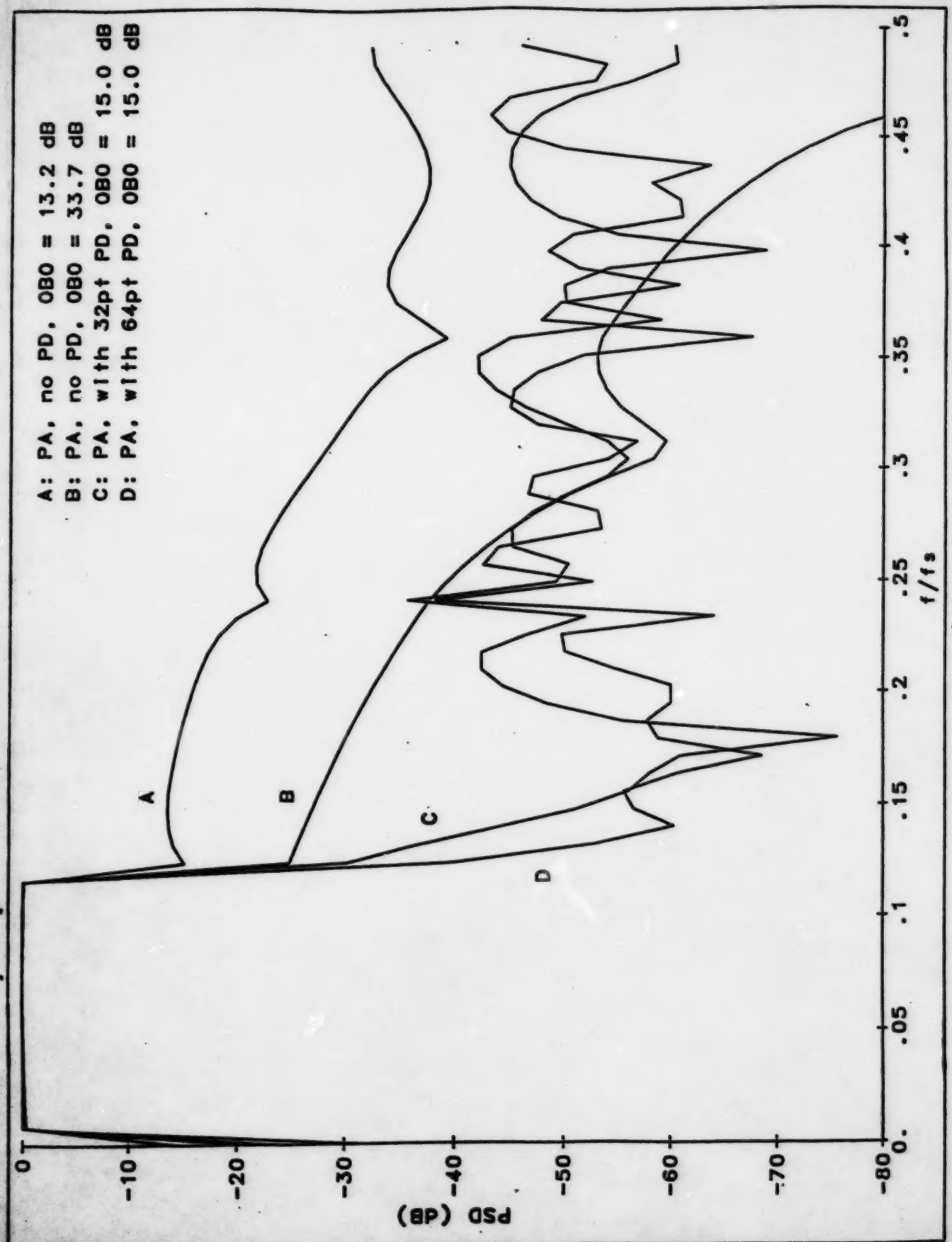

Figure 8: Noise Loading Test, Showing Effect of Predistorter

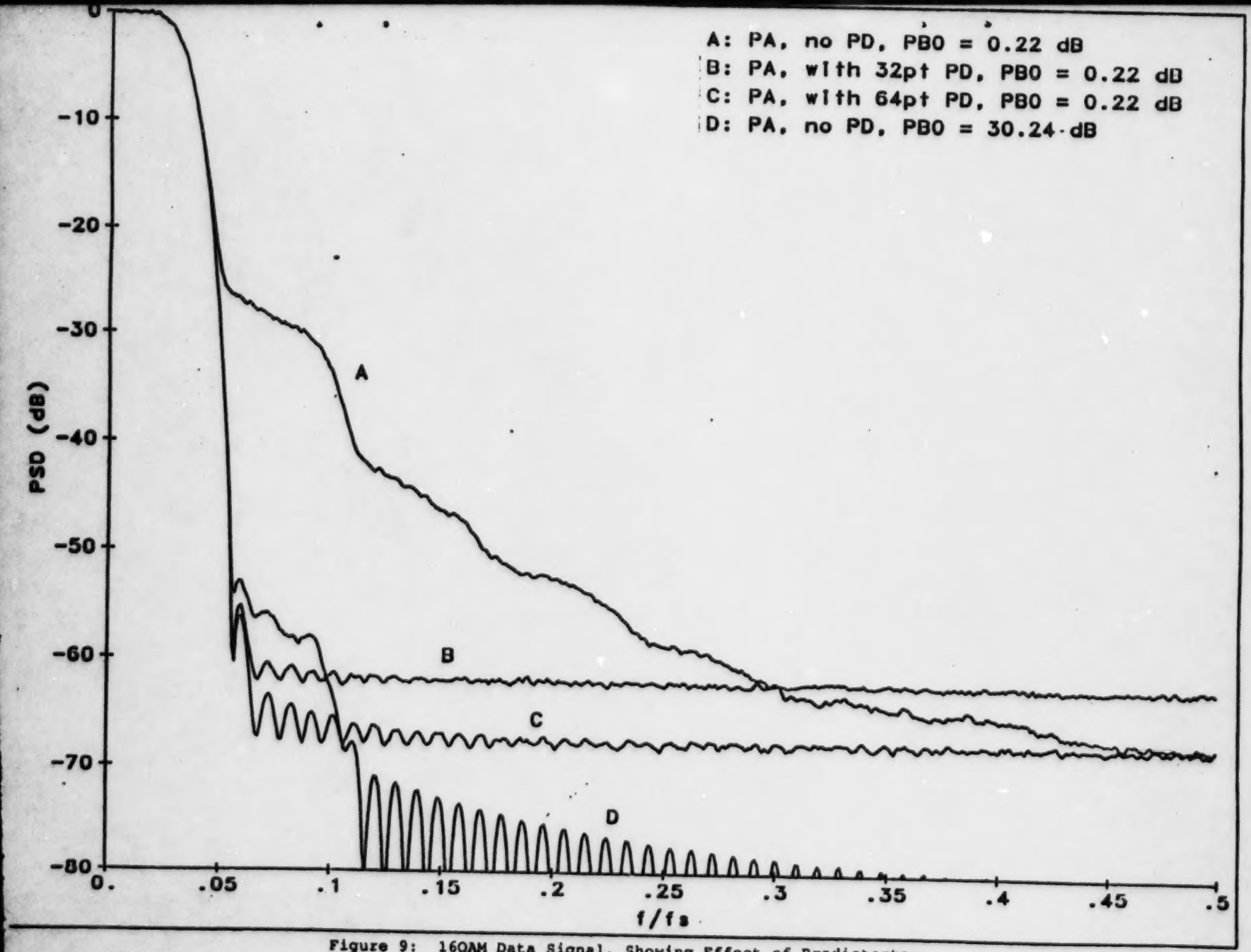

Figure 9: 16QAM Data Signal, Showing Effect of Predistorter

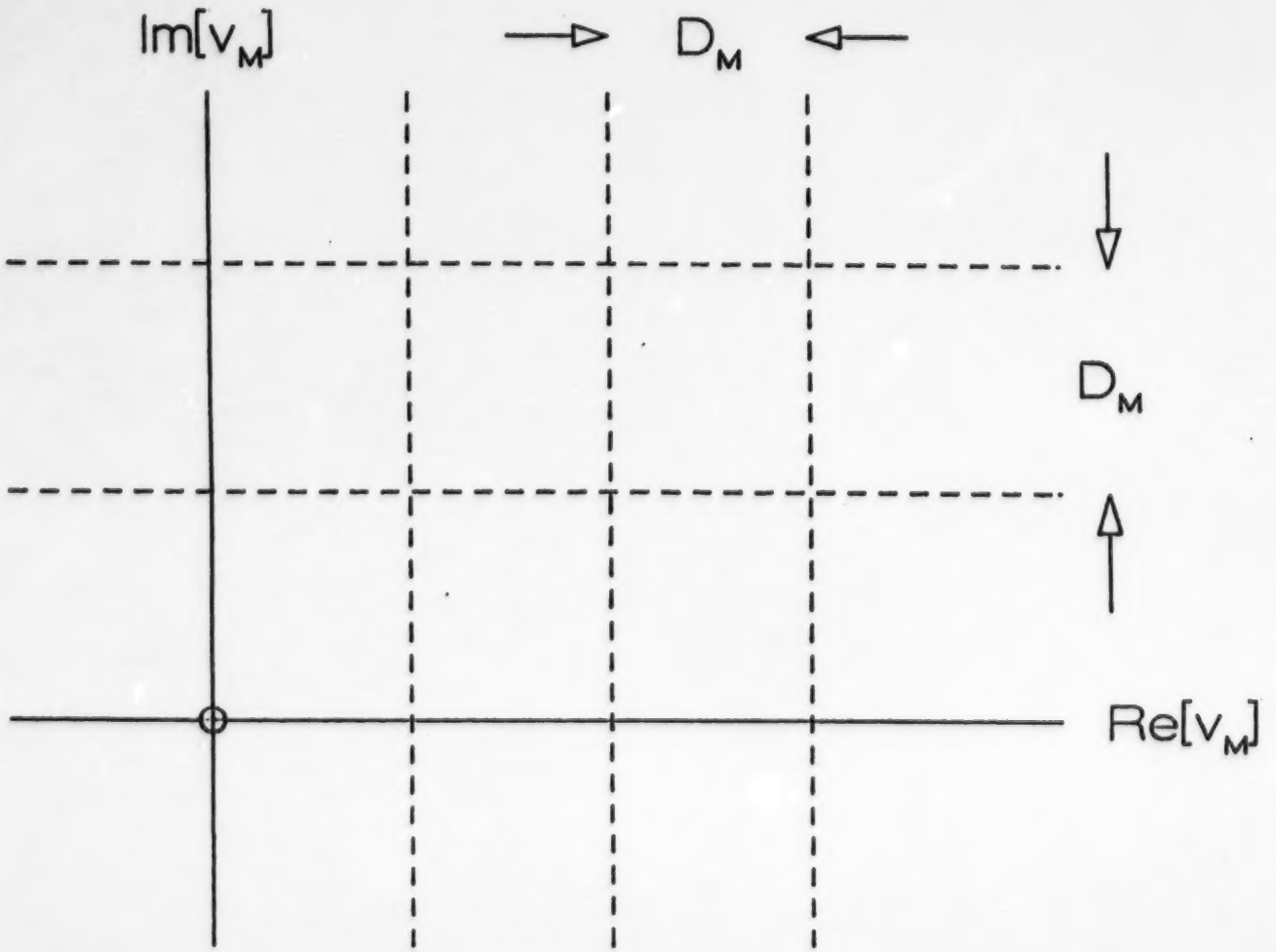

Figure 10: Input Quantization in the Mapping Predistorter

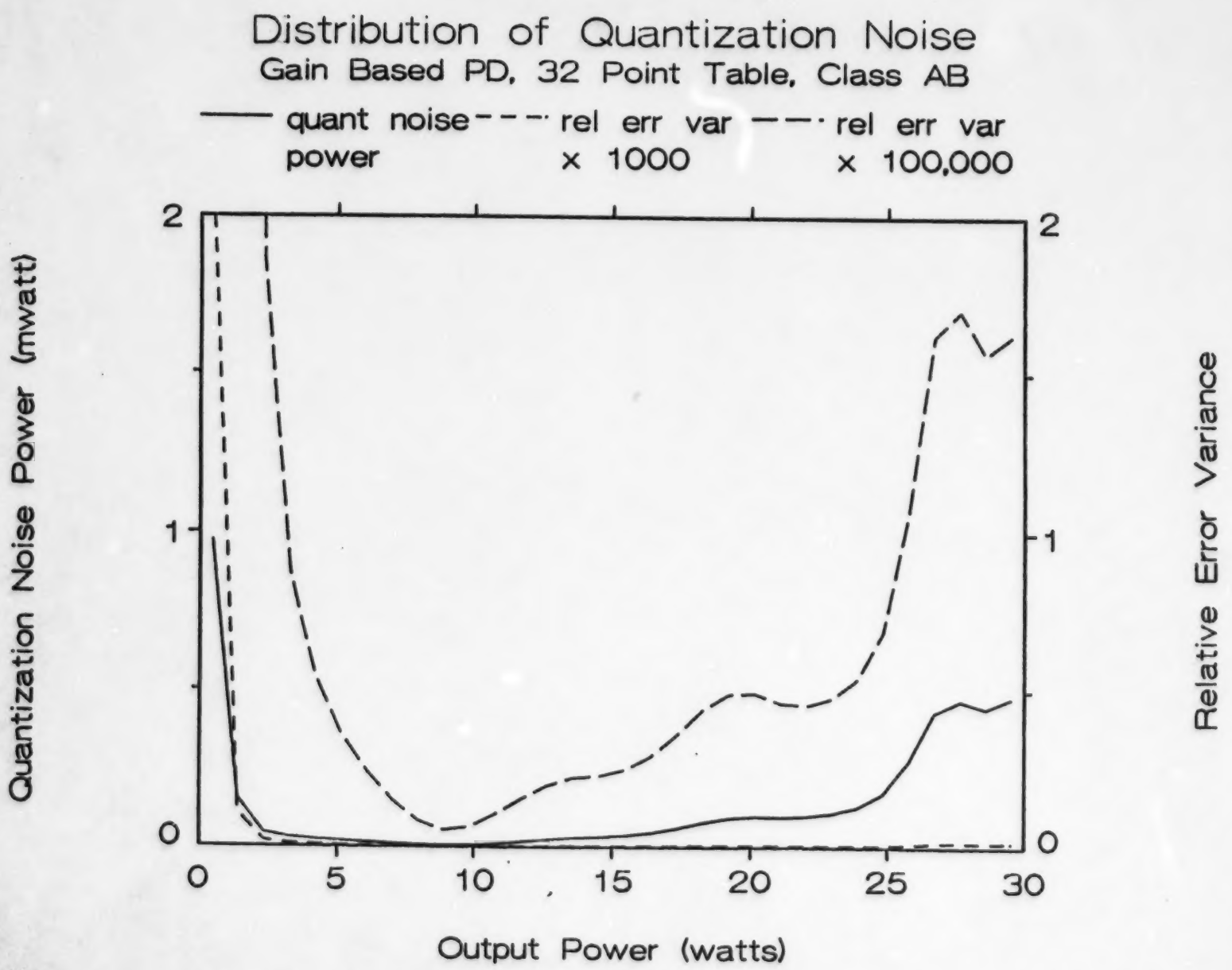

Figure 11: Distribution of Intermod Power With Instantaneous Output Level Using Gain Based Predistorter

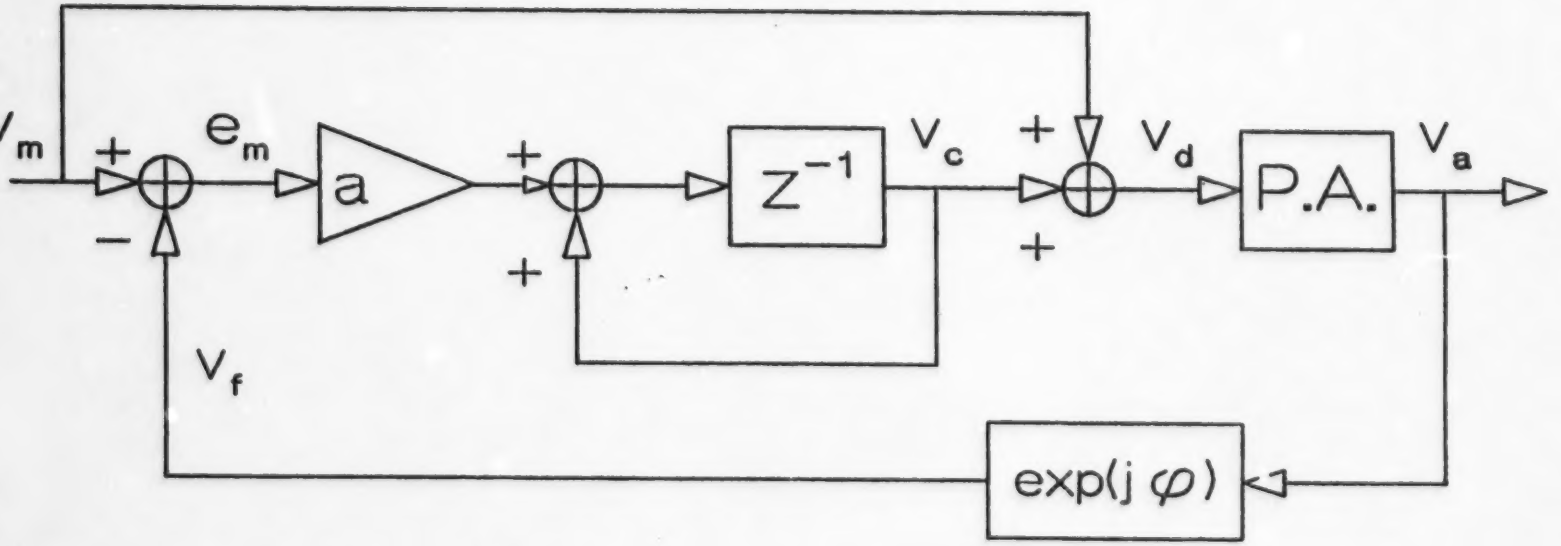

Figure 12: A Linear Adaptation Update Structure

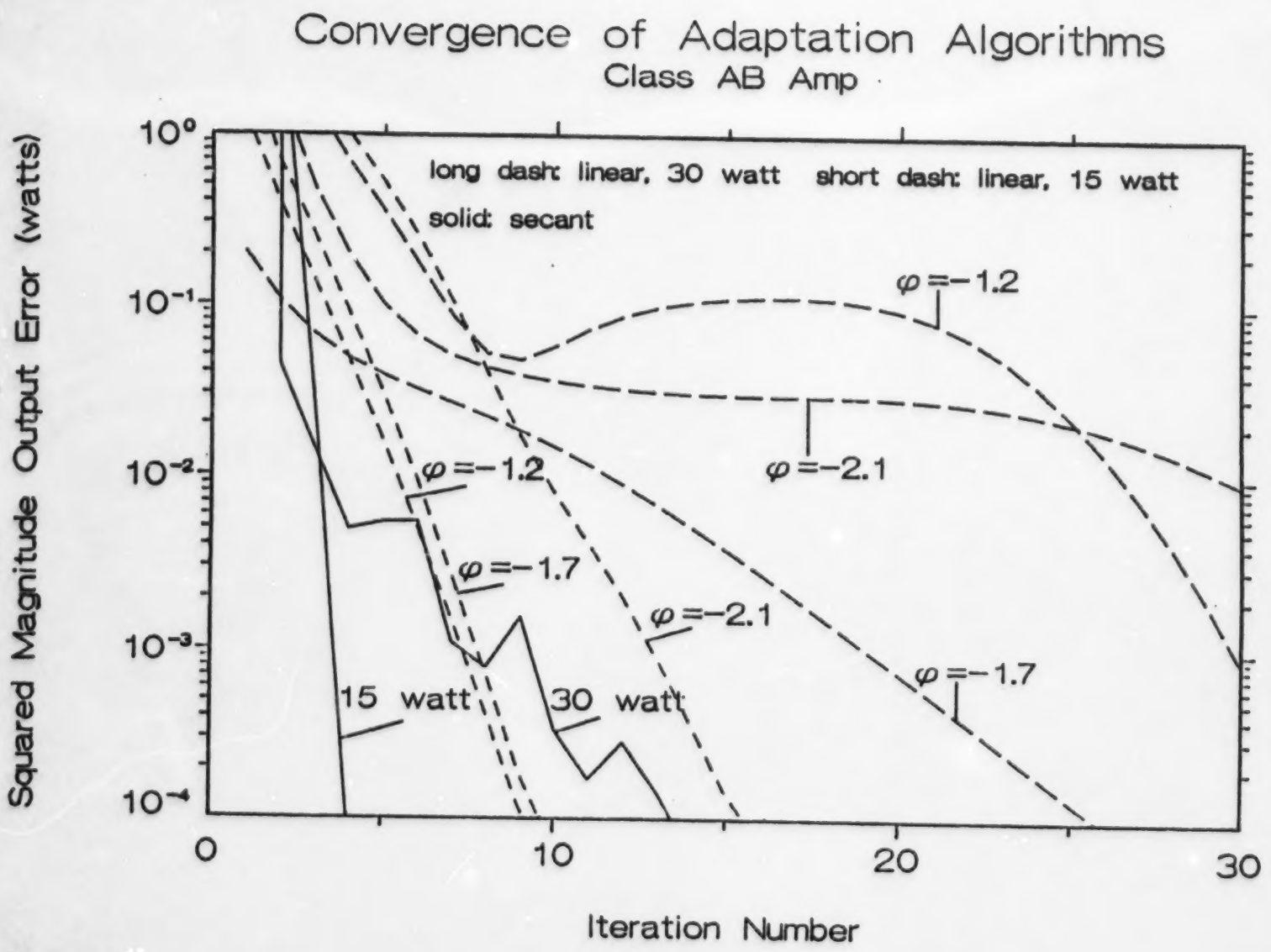

Figure 13: Convergence Behaviour of Linear and Secant Update Algorithms

Centre for Systems Science

• Simon Fraser University established the **Centre for Systems Science** as a provincial centre for advanced study. The objective of CSS is twofold: coordinate and combine advanced systems research at SFU, and encourage application of research results within industry in order to further the technological development of British Columbia.

• CSS has assembled three research groups to streamline its efforts. The orientation of these groups is as follows: intelligent systems, microelectronics, and computer and communications systems. Because of shared research interests, this latter group merges with LCCR (**Laboratory for Computer and Communications Research**) for the publication of technical reports.

• CSS's three groups involve over 80 skilled faculty researchers principally from SFU's Schools of Engineering Science and Computing Science, but also draw substantially on Mathematics, Physics, Geography, Philosophy, Linguistics, Psychology, Communication, and Kinesiology.